Niveau(x) d'enseignement : 0 Numéro EDUQ : 8970845
Centre d'intérêt : 540

Guide d'élaboration d'instruments de mesure / Direction du développement de l'évaluation, Direction générale de l'évaluation et des ressources didactiques. -- [Québec] : Ministère de l'éducation, c1990.
41, [5] p. : ill. ; 28 cm.

Code : 16-7290. Bibliogr. : p. [43-46].

ISBN 2-550-15175-5

Vedette(s)-secondaire(s) : Québec (Province). Direction du développement de l'évaluation

Descripteur(s) : Evaluation des apprentissages • Instrument de mesure • Elaboration d'instruments • Guide

Identificateur(s) de sujet : Apprentissage cognitif

Ce guide présente les étapes à franchir et les règles à respecter lors de l'élaboration d'instruments de mesure permettant d'évaluer les apprentissages cognitifs. Le premier chapitre présente le concept de définition de domaine et précise son rôle dans l'élaboration d'un instrument de mesure. Il traite également de la nécessité de sélectionner judicieusement les types et le nombre d'items. Le deuxième chapitre expose les règles décimologiques à respecter lors de la rédaction des items, après avoir défini les différents types d'items et décrit les avantages et les limites de chacun. Un dernier chapitre est consacré à l'assemblage et à l'administration de l'épreuve.

Type de document : d

8970845

GUIDE D'ÉLABORATION D'INSTRUMENTS DE MESURE

Québec

GUIDE D'ÉLABORATION D'INSTRUMENTS DE MESURE

Direction du développement de l'évaluation
Direction générale de l'évaluation et des ressources didactiques

© Gouvernement du Québec
Ministère de l'Éducation, 1990

ISBN 2-550-15175-5

Dépôt légal — premier trimestre 1990
Bibliothèque nationale du Québec

SOMMAIRE

INTRODUCTION

CHAPITRE 1. SPÉCIFIER LE CONTENU DE L'ÉPREUVE

1.1 Déterminer le but de l'épreuve

1.2 Utiliser la définition de domaine

1.2.1 La présentation de la définition de domaine

1.2.1.1 Les orientations du programme et les principes à respecter en évaluation
1.2.1.2 La description des éléments du programme
1.2.1.3 Le tableau des dimensions

1.2.2 L'utilisation de la définition de domaine en cours d'année

1.3 Choisir les types d'items

1.4 Déterminer le nombre d'items

CHAPITRE 2. RÉDIGER LES ITEMS

2.1 Respecter les règles docimologiques

2.1.1 Les parties d'un item

2.1.2 Les règles de rédaction communes à tous les types d'items

2.1.3 Les items à réponse choisie

2.1.3.1 Les items à choix multiple
A. Définition
B. Avantages et limites
C. Règles de rédaction

2.1.3.2 Les items vrai ou faux
A. Définition
B. Avantages et limites
C. Règles de rédaction

2.1.3.3 Les items de type appariement
A. Définition
B. Avantages et limites
C. Règles de rédaction

2.1.4 Les items à réponse construite

2.1.4.1 Les items à réponse courte
A. Définition
B. Avantages et limites
C. Règles de rédaction

2.1.4.2 Les items à réponse élaborée
A. Définition
B. Avantages et limites
C. Règles de rédaction
D. Méthodes de correction

2.1.4.3 Les grilles d'observation
A. Définition
B. Avantages et limites
C. Règles de construction

2.2 Réviser et corriger les items

CHAPITRE 3. ASSEMBLER ET ADMINISTRER L'ÉPREUVE

3.1 Assembler l'épreuve

3.1.1 Le cahier de l'élève

3.1.1.1 L'ordre des items
3.1.1.2 Les directives pour l'élève
3.1.1.3 Le cahier ou la feuille de réponses

3.1.2 Le guide de l'enseignant ou de l'enseignante

3.2 Administrer l'épreuve

3.2.1 Mettre à l'essai et corriger l'épreuve

3.2.2 Expérimenter l'épreuve

BIBLIOGRAPHIE

INTRODUCTION

La finalité de l'évaluation des apprentissages est la prise de décisions justes et équitables pour tous les élèves.

Quelle que soit l'intention d'évaluation, à savoir la sanction des études ou la régulation des apprentissages, la démarche reste la même. Elle se déroule en trois étapes : la mesure, le jugement et la décision.

La première étape de la démarche d'évaluation, celle de la mesure, consiste à recueillir des informations ou des données, à les traiter puis à les interpréter. Pour assurer la valeur des décisions prises à la suite de l'interprétation des données recueillies, il importe que la collecte des informations se fasse avec beaucoup de rigueur.

La cueillette d'informations peut se réaliser entre autres par l'utilisation d'instruments de mesure. Le présent guide traite de ce moyen de recueillir de l'information, sans qu'on y aborde toutefois les aspects de l'organisation et de l'interprétation des données. Il présente essentiellement les étapes à franchir et les règles à respecter lors de l'élaboration d'instruments de mesure permettant d'évaluer les apprentissages cognitifs.

Dans le guide, chacune de ces étapes est explicitée.

Ainsi, le chapitre 1 présente le concept de définition de domaine ainsi que son rôle dans l'élaboration d'un instrument de mesure. Il traite également de la nécessité de sélectionner judicieusement les types et le nombre d'items.

Au chapitre 2, on expose les règles docimologiques à respecter lors de la rédaction des items, après avoir défini les différents types d'items et décrit les avantages et les limites de chacun.

Le chapitre 3 est consacré à l'assemblage et à l'administration de l'épreuve.

Ce guide s'adresse à toute personne appelée à superviser, coordonner ou participer à la conception d'un instrument de mesure des apprentissages en formation générale.

REMERCIEMENTS

La Direction du développement de l'évaluation du ministère de l'Éducation tient à remercier Pierrette Jalbert, Joanne Munn et Diane Provençal, de la Banque d'instruments de mesure (BIM) de la Société GRICS, qui ont conçu et réalisé le présent guide, et à souligner la contribution des personnes suivantes qui l'ont lu, analysé et commenté tant au plan des idées que de la forme :

Lisette Côté-Bernier	Ministère de l'Éducation
Pierre Bérubé	Commission scolaire régionale de Chambly
Monique Brunelle	Ministère de l'Éducation
Johanne Coulombe	Banque d'instruments de mesure (BIM), Société GRICS
Chantale Deschamps	Commission scolaire de Greenfield Park
Marie-Sylvie Descôteaux	Ministère de l'Éducation
Michel Doré	Ministère de l'Éducation
Michel Fréchette	Banque d'instruments de mesure (BIM), Société GRICS
Claude Girard	Commission scolaire de Beauport
Annie Roberge	Commission scolaire des Mille-Îles

CHAPITRE 1

SPÉCIFIER LE CONTENU DE L'ÉPREUVE

1.1 DÉTERMINER LE BUT DE L'ÉPREUVE

Pour déterminer le but de l'épreuve, il faut prendre en considération la nature de la décision qui sera prise à partir des résultats obtenus.

Si l'on désire, à la fin d'une étape ou au terme d'une année scolaire, dresser le portrait des apprentissages réalisés en vue d'un rapport périodique ou de la sanction des études, il est souhaitable d'élaborer l'épreuve en respectant une démarche structurée comme celle qui est proposée dans ce guide.

Par contre, dans un contexte d'évaluation continue en classe, afin d'ajuster l'enseignement, il y a lieu d'avoir recours à d'autres techniques comme l'observation directe, l'objectivation, l'interrogation orale, etc. Ces techniques sont décrites, entre autres, dans le Guide d'évaluation en classe[1] de chacun des programmes d'études concerné.

1.2 UTILISER LA DÉFINITION DE DOMAINE

Lors de la conception d'une épreuve portant sur l'ensemble d'un programme d'études, il faut d'abord et avant tout s'assurer que les items qui composent l'épreuve permettent de mesurer l'ensemble des apprentissages prescrits par ce programme. Autrement dit, il est fondamental de s'assurer de la congruence entre les apprentissages prévus dans le programme et les items de l'épreuve élaborée à partir de ce programme.

Plusieurs méthodes permettent d'assurer la représentativité des items d'une épreuve en regard d'un programme d'études. Les plus utilisées actuellement au Québec sont la définition de domaine[2] et le tableau de spécifications élaboré à partir de la taxonomie de Bloom.

Dans ce guide, on privilégie la première approche, soit l'utilisation de la définition de domaine. C'est aussi celle qui est utilisée pour l'élaboration des épreuves en formation générale (épreuves d'appoint, épreuves de sanction des études, épreuves d'évaluation de programme).

1.2.1 La présentation de la définition de domaine

La définition de domaine expose l'ensemble des éléments essentiels et représentatifs d'un programme d'études, ensemble dont il faut tenir compte lors de l'élaboration d'instruments de mesure des apprentissages visés par ce programme.

---

[1] Les guides d'évaluation en classe (G.E.C.) ont été produits dans le but d'apporter un soutien à l'enseignant ou à l'enseignante dans sa démarche d'évaluation formative. Chacun des guides rappelle brièvement les orientations du programme concerné et donne des exemples d'évaluation formative. Le répertoire des publications de la Direction du développement de l'évaluation dresse la liste des G.E.C. disponibles.

[2] Auger, R., Fréchette, M. La définition de domaine, une étape essentielle dans l'élaboration d'un instrument de mesure, *«Mesure et évaluation en éducation»*, vol. 10, n° 4, hiver 1988, p. 5 à 22.

La définition de domaine est présentée dans un document comprenant généralement trois parties :

- Les orientations du programme et les principes à respecter en évaluation
- La description des éléments du programme d'études
- Le tableau des dimensions

En se référant à ce document, on s'assure de construire une épreuve en conformité avec le programme d'études.

1.2.1.1 Les orientations du programme et les principes à respecter en évaluation

Dans cette partie de la définition de domaine, on rappelle sommairement les orientations, les objectifs généraux et les buts d'un programme d'études donné. On y présente aussi les principes à respecter lors de l'élaboration d'une épreuve. Ces principes, qui tiennent compte des orientations du programme, précisent le contexte dans lequel doit se réaliser l'évaluation des apprentissages.

Par exemple :

- Le programme de mathématique au 2e cycle du secondaire vise à développer des habiletés et à mettre en lumière le lien qui existe entre la mathématique et la réalité. On devra tenir compte de ces orientations en vérifiant si l'élève a effectivement développé des habiletés, et cela, autant que possible, à travers des mises en situation variées, réalistes, proches du vécu de l'élève.

- Le programme de français est fondé sur l'approche communicative. Par conséquent, il faudra placer l'élève dans des situations signifiantes où il devra exercer ses habiletés à communiquer soit par écrit, soit oralement.

- En histoire du Québec et du Canada, le programme vise entre autres à développer une compréhension globale de l'histoire. Ainsi, on évaluera la capacité de l'élève à analyser des réalités historiques selon plusieurs aspects de société.

1.2.1.2 La description des éléments du programme d'études

Cette partie décrit les éléments représentatifs d'un programme d'études. On y identifie, d'une part, les contenus (thèmes, modules, unités, concepts, etc.) et, d'autre part, les habiletés que les élèves sont appelés à développer et à exercer pour traiter ces contenus.

Par exemple :

En mathématique au primaire, l'élève doit être capable de

HABILETÉS	THÈMES
Structurer Mathématiser Opérer Analyser/Synthétiser	Les nombres La géométrie La mesure

En géographie du Québec et du Canada, l'élève doit être capable de

HABILETÉS	CONCEPTS
Décrire	L'espace
Analyser	Les ressources
Synthétiser	Les activités économiques

Afin de s'assurer d'une compréhension univoque de ces éléments (contenus et habiletés), on fournit, dans cette partie de la définition de domaine, une liste des principaux contenus.

Par exemple :

En mathématique au primaire, on rappelle les principales notions ou les principaux concepts qui sont regroupés sous le thème des nombres, de la géométrie et de la mesure.

Nombres naturels :
- nombre comme propriété
- numération et valeur de positions
- lecture et écriture
- etc.

Géométrie et mesure :
- régions
- solides
- symétrie
- etc.

Par ailleurs, on donne aussi une définition des habiletés. Chaque habileté est définie en tenant compte à la fois des orientations, des principes directeurs et des objectifs du programme d'études, et à chacune d'elles correspond une liste de manifestations ou de comportements observables.

Par exemple :

En mathématique au primaire, l'habileté à mathématiser signifie que l'élève est capable de traduire une situation donnée par un modèle mathématique : arithmétique, algébrique ou graphique. Cette habileté se manifeste lorsque l'élève doit illustrer, représenter, schématiser, symboliser, traduire ou transposer une situation donnée.

1.2.1.3. Le tableau des dimensions

Dans cette partie de la définition de domaine, on présente, sous forme de tableau à double entrée, les contenus et les habiletés. Dans ce tableau, on appelle «dimension» l'intersection entre une habileté et un contenu. Chaque dimension précise l'objet de mesure à partir duquel les items seront rédigés.

À la page suivante (tableau 1), on présente, à titre d'exemple, le tableau des dimensions de la définition de domaine du programme de mathématique au 2e cycle du primaire.

Tableau 1

TABLEAU DES DIMENSIONS
Mathématique, primaire, 2e cycle

THÈMES / HABILETÉS	Nombres naturels Entiers relatifs 40 %	Nombres rationnels 25 %	Géométrie Mesure 35 %
STRUCTURER 25 %	1	4	7
MATHÉMATISER 20 %	2	5	8
OPÉRER 25 %	3	6	9
ANALYSER OU SYNTHÉTISER 30 %	10		

Légende: Les nombres 1 à 10 identifient les dimensions.

Dans ce tableau des dimensions, il est à remarquer que l'on précise aussi l'importance relative de chaque habileté et de chaque contenu. Cette importance relative a une incidence directe sur l'importance qu'aura chaque item par rapport à l'ensemble des items de l'épreuve. Ainsi, dans cet exemple emprunté à la mathématique, on constate que l'importance accordée aux nombres rationnels est de 25 %; conséquemment, les items portant sur les nombres rationnels devront représenter 25 % de l'épreuve. De la même façon, puisque l'importance relative accordée à la synthèse y est de 30 %, celle accordée aux items de synthèse dans l'épreuve devra également être de 30 %.

Après la lecture de cette partie de la définition de domaine, on sait quelle doit être la répartition des items de l'épreuve pour qu'on représente adéquatement l'ensemble des apprentissages visés par le programme d'études.

L'utilisation de la définition de domaine lors de l'élaboration d'une épreuve permet donc de s'assurer du lien entre l'épreuve et le programme d'études. En effet, si l'on rédige les items selon les orientations du programme, en couvrant l'ensemble des dimensions tout en respectant l'importance relative indiquée dans le tableau des dimensions, on a la certitude que l'épreuve est le portrait fidèle du programme d'études.

La définition de domaine constitue en quelque sorte un ensemble de données à partir duquel plusieurs épreuves, portant sur un même programme d'études, peuvent être produites.

1.2.2 L'utilisation de la définition de domaine en cours d'année

Comme nous l'avons mentionné précédemment, c'est particulièrement lorsque l'on construit une épreuve pour mesurer l'ensemble des apprentissages prescrits dans un programme d'études qu'il y a lieu d'utiliser la définition de domaine, celle-ci ayant été conçue pour des fins d'évaluation sommative. Cependant, si l'on désire mesurer les apprentissages visés dans une partie du programme (un module, une unité, un volet...), on peut aussi recourir à la définition de domaine. Dans ce cas, il faut identifier les dimensions à retenir, et réajuster l'importance relative accordée aux contenus et aux habiletés.

En utilisant la définition de domaine en cours d'année, on a l'avantage de tracer le portrait de l'élève à partir du même cadre de travail que celui qui sera utilisé en fin d'année.

Exemple 1

On désire vérifier la maîtrise de la partie statistique en mathématique, première secondaire. Pour cela, il faut retenir trois des neuf dimensions de la définition de domaine, soit les dimensions 7, 8 et 9 (tableau 2a, ci-dessous). Étant donné que la dimension «opérer en statistique» n'existe pas, seulement trois des quatre habiletés seront mesurées dans l'épreuve. Il est donc nécessaire de modifier l'importance relative de chacune en répartissant le 25 % initialement accordé à l'habileté à opérer.

Tableau 2a

TABLEAU DES DIMENSIONS
Mathématique, 1re secondaire

THÈMES / HABILETÉS	NOMBRES 55 %	GÉOMÉTRIE ET MESURE 35 %	STATISTIQUE 10 %
STRUCTURER 25 %	1	4	7
MATHÉMATISER 20 %	2	5	8
OPÉRER 25 %	3	6	
ANALYSER OU SYNTHÉTISER 30 %	9		

Légende *: Les nombres 1 à 9 identifient les dimensions.*

À la suite de ces modifications, on obtiendra le tableau suivant :

Tableau 2b

TABLEAU DES DIMENSIONS RETENUES
Mathématique, 1re secondaire (statistique)

THÈME / HABILETÉS	STATISTIQUE
STRUCTURER 50 %	Lecture de divers diagrammes (6.1)
MATHÉMATISER 40 %	Construction de tableaux de données statistiques (6.2)
OPÉRER	//////////
ANALYSER OU SYNTHÉTISER 10 %	Résolution de problèmes (6.1 - 6.2)

Légende : Les numéros entre parenthèses font référence aux numéros des objectifs terminaux du programme d'études.

Exemple 2

En mathématique au primaire, les différents thèmes sont enseignés toute l'année. Ce sont les notions rattachées à ces thèmes qui varient d'une étape à l'autre. Par conséquent, il est possible que certaines dimensions ne soient pas mesurées bien que l'ensemble des habiletés et l'ensemble des thèmes le soient. Ainsi, le tableau 1 (page 4) sera modifié car il faut réajuster l'importance relative des habiletés et des thèmes. Le tableau 3 illustre une situation où les dimensions 2, 6 et 9 ne sont pas mesurées.

Tableau 3

TABLEAU DES DIMENSIONS RETENUES
Mathématique, primaire, 2e cycle

THÈMES / HABILETÉS	Nombres naturels Entiers relatifs 40 %	Nombres rationnels 25 %	Géométrie Mesure 35 %
STRUCTURER 40 %	1 . numération et valeur de position (1) . relation d'ordre (2-7)	4 . sens de la fraction (8) . relation d'ordre (9-10)	7 . solides (16) . figure à deux dimensions (17)
MATHÉMATISER 20 %	2 //////////////	5 . représentation d'une situation (12-13)	8 . représentation d'une situation : périmètre, circonférence, volume (26)
OPÉRER 10 %	3 . addition (4) . soustraction (4)	6 //////////////////	9 //////////////////////
ANALYSER OU SYNTHÉTISER 30 %	10 résolution de problèmes (1-2-4-7-8-9-10-12-15-16-17-26)		

Légende : ////// *indique que ces dimensions ne sont pas mesurées.*

Les numéros entre parenthèses font référence aux numéros des objectifs terminaux du programme d'études.

Enfin, dans les cas où l'on souhaite bâtir une épreuve pour ne mesurer qu'un seul objectif, il n'est pas nécessaire de recourir à la définition de domaine. Il peut être suffisant de s'assurer de la congruence entre les items et l'objectif du programme d'études. Pour ce faire, il faut se demander si chaque item mesure bien l'objectif visé.

1.3 CHOISIR LES TYPES D'ITEMS

Il existe plusieurs types d'items. Au chapitre 2, on présente la définition, les avantages et les limites ainsi que les règles de rédaction de chacun d'eux. Avant de choisir un type d'item, on peut d'abord consulter ce chapitre.

Plusieurs facteurs interviennent au moment du choix d'un type d'item. Il faut considérer le but de l'épreuve, la dimension à mesurer et la réponse fournie par l'élève. Avant d'arrêter définitivement le choix du type d'item, il faut aussi tenir compte des deux facteurs suivants : la correction et le temps d'administration. Dans les lignes qui suivent, on examine séparément ces facteurs. Le choix d'un type d'item doit s'effectuer en combinant toutes les informations relatives à tous ces facteurs.

Le but de l'épreuve

Lorsque les épreuves sont utilisées pour la sanction des études ou pour le classement des élèves, il est très important de s'assurer de la fidélité des résultats, notamment par l'uniformité de la correction. Par conséquent, les items à réponse choisie doivent être privilégiés chaque fois que cela est possible. En classe, comme l'enseignant ou l'enseignante corrige les copies de ses propres élèves, le problème relié à l'uniformité de la correction est moindre. Par conséquent, on peut utiliser les items à réponse construite. Ces derniers sont d'autant plus intéressants dans ce contexte car ils fournissent des informations relatives à la démarche de l'élève.

De façon générale, tout au cours de l'année scolaire, il est important d'habiliter les élèves à répondre à tous les types d'items.

La dimension

Pour chaque dimension du tableau des dimensions, il faut déterminer le type d'item qui correspond le mieux à l'habileté et au contenu visés. En d'autres mots, il faut s'assurer de la congruence entre les items et les dimensions.

Dans certains cas, il sera tout aussi indiqué de recourir à l'un ou à l'autre des types d'items. Par contre, il peut arriver que pour mesurer une dimension donnée, les items à réponse construite soient les seuls qui conviennent, par exemple.

Il peut être nécessaire lors de cette étape d'étudier plus en détail le contenu[1] et l'habileté visés dans une dimension. À ce moment, il faut examiner plus particulièrement les éléments décrits dans la deuxième partie de la définition de domaine, c'est-à-dire les manifestations ou les comportements observables pour chaque habileté ainsi que la liste des contenus.

La réponse

Les caractéristiques de la réponse attendue influent sur le choix des types d'items. Il peut s'agir d'une réponse courte comme un mot, une lettre ou un symbole. Ce peut être aussi une réponse élaborée comme dans le cas d'un graphique, d'un schéma, d'un texte, etc.

De plus, il faut tenir compte du fait que la réponse peut être unique, c'est-à-dire la même pour tous les élèves, ou encore, qu'elle peut être différente pour chacun (par exemple: productions orales et écrites, problèmes à résoudre).

Pour sélectionner le type d'items en tenant compte à la fois de la dimension et de la réponse de l'élève, on peut s'aider des questions et des informations présentées ci-dessous.

1. À partir d'un stimulus semblable à ceux présentés en classe, l'élève doit-il se rappeler les éléments appris?

   Par exemple, l'élève doit-il :

   . répéter une définition ou énoncer une loi?
   . localiser une structure anatomique sur un schéma?
   . situer dans le temps un événement historique?
   . situer sur une carte géographique un bassin hydrographique?
   . etc.

---

[1] Selon chacun des programmes d'études, les contenus peuvent être des faits (tels des noms de personnages), des lieux, des mots de vocabulaire, des définitions. Il peut s'agir aussi de concepts, de lois, de relations, ou encore, de structures plus complexes comme des cartes géographiques, des textes, etc.

Si l'on répond oui à cette question, il est recommandé d'utiliser les items à réponse choisie, ceux à choix multiple ou ceux de type appariement parce que généralement, la réponse à ces items est identique pour tous les élèves et elle est courte (un mot, un symbole, une lettre, etc.). On peut utiliser aussi des items à réponse construite ou à réponse courte. Les items vrai ou faux ne sont pas recommandés dans ce cas.

2. A partir d'un stimulus différent mais ayant les mêmes caractéristiques que ceux présentés en classe, l'élève doit-il utiliser des éléments appris?

   Par exemple, l'élève doit-il :

   . classer des observations, en sciences?
   . appliquer des lois, des règles ou des théorèmes?
   . analyser un temps fort historique?
   . etc.

Il est difficile de formuler ici une recommandation qui soit valable pour tous les programmes d'études. Les habiletés et les contenus prescrits dans chaque programme sont trop diversifiés. Toutefois, à titre d'exemple, on peut examiner les deux situations présentées ci-dessous :

a) Le contenu à traiter est restreint, et la réponse est courte et unique (par exemple, en mathématique, lorsque l'élève doit identifier tous les triangles isocèles dans une figure).

b) Le contenu à traiter est large, et la réponse est élaborée et unique (par exemple, en mathématique, lorsque l'élève doit effectuer une transformation géométrique).

Dans le premier cas, on privilégie les items à réponse choisie tandis que dans le deuxième, on choisit les items à réponse construite.

3. A partir d'un stimulus complètement différent de ceux présentés en classe, l'élève doit-il combiner plusieurs éléments?

   Par exemple, l'élève doit-il :

   . résoudre un problème en mathématique ou en éducation économique?
   . rédiger un texte en français ou en anglais?
   . etc.

Si l'on répond oui à cette question, il est recommandé d'utiliser les items à réponse construite et particulièrement ceux à réponse élaborée. Généralement, le contenu abordé est large et la réponse est différente pour chaque élève. Par exemple, en français, pour rédiger un texte, l'élève doit réunir en un tout cohérent plusieurs éléments (fonctionnement de la langue, fonctionnement du discours, etc.) abordés sous divers aspects en classe. Les textes ainsi rédigés sont des productions originales.

On peut aussi utiliser certains items à réponse choisie, comme ceux à choix multiple ou les items vrai ou faux, ainsi que les items à réponse courte si la réponse de l'élève est courte et unique. Par exemple, en mathématique ou en sciences, on peut recourir à ces types d'items si l'on ne corrige que le résultat d'un problème.

Les conditions relatives à la correction

Dans le cas où il est tout aussi possible d'utiliser les items à réponse choisie que les items à réponse construite, il faut tenir compte des conditions relatives à la correction. Par exemple, lorsque le délai est très court entre l'administration de l'épreuve et la consignation des résultats, il est préférable d'avoir recours aux items à réponse choisie. Il en sera de même s'il n'est pas possible d'assurer l'uniformité de la correction des items à réponse construite, particulièrement pour les épreuves servant à des fins sommatives.

Le temps d'administration

Le temps alloué à l'administration de l'épreuve affecte le choix des types d'items. Une fois cette durée déterminée (45, 50, 60 minutes...), il faut tenir compte des caractéristiques des items; par exemple, les élèves répondent plus rapidement à des items à réponse choisie qu'à des items à réponse construite.

LA DÉTERMINER LE NOMBRE D'ITEMS

En fonction de la durée prévue pour l'administration de l'épreuve et des types d'items retenus pour chacune des dimensions, il faut déterminer le nombre total d'items à rédiger en tenant compte de l'importance relative indiquée au tableau des dimensions ou réajustée selon les circonstances. Une fois ce nombre déterminé, on indique à l'intérieur de chacune des dimensions le ou les types d'items retenus de même que le nombre d'items de chaque type.

Dans les deux exemples de la page suivante, on reprend deux tableaux (tableaux 2b et 3) en les modifiant pour illustrer deux adaptations de la définition de domaine en mathématique.

Le premier exemple présente le tableau des dimensions d'une épreuve de 50 minutes en statistique en première secondaire, contenant six items (tableau 4).

Tableau 4

TABLEAU INDIQUANT LA RÉPARTITION DES ITEMS
Mathématique, 1re secondaire (statistique)

THÈME / HABILETÉS	STATISTIQUE
STRUCTURER 50 %	3 items à réponse courte
MATHÉMATISER 40 %	2 items à réponse élaborée
OPÉRER	//////////////////////////////////////
ANALYSER OU SYNTHÉTISER 10 %	1 item à réponse élaborée

Le deuxième exemple montre le tableau des dimensions d'une épreuve pour le 2e cycle du primaire, constituée de 20 items (tableau 5).

Tableau 5

TABLEAU INDIQUANT LA RÉPARTITION DES ITEMS
Mathématique, primaire, 2e cycle

THÈMES / HABILETÉS	Nombres naturels Entiers relatifs 40 %	Nombres rationnels 25 %	Géométrie Mesure 35 %
STRUCTURER 40 %	3 items à réponse courte	2 items à réponse courte	3 items à réponse courte
MATHÉMATISER 20 %		2 items à réponse courte	2 items à réponse courte
OPÉRER 10 %	2 items à réponse courte		
ANALYSER OU SYNTHÉTISER 35 %	6 items à réponse élaborée		

CHAPITRE 2

RÉDIGER LES ITEMS

2.1 RESPECTER LES RÈGLES DOCIMOLOGIQUES

2.1.1 Les parties d'un item

Un item est un élément d'un instrument de mesure. Ce peut être un problème à résoudre, une question ou une tâche plus ou moins complexe.

Il existe plusieurs types d'items, qui se répartissent en deux grandes catégories: les items à réponse choisie et les items à réponse construite.

LES ITEMS À RÉPONSE CHOISIE

Les items à réponse choisie sont ceux pour lesquels des réponses sont formulées à l'avance; l'élève doit choisir parmi ces réponses celle qu'il croit être la réponse correcte ou admise. Font partie de cette catégorie les items à choix multiple, les items vrai ou faux, les items de type appariement.

LES ITEMS À RÉPONSE CONSTRUITE

Les items à réponse construite sont ceux dont la réponse doit être formulée par l'élève. Dans cette catégorie, on retrouve les items à réponse courte et les items à réponse élaborée. Il arrive souvent que ces derniers soient corrigés à l'aide de grilles d'observation.

Un item est composé des parties suivantes:

1) *Le stimulus*

Le stimulus présente la tâche que l'élève doit réaliser. Il est constitué essentiellement d'une question ou d'une consigne. Dans certains cas, une mise en situation précède la question ou la consigne.

2) *La réponse*

- Pour les items à réponse choisie, la réponse est choisie par l'élève parmi un choix de réponses proposées.
- Pour les items à réponse construite, la réponse est formulée par l'élève lui-même.

3) *La clé de correction*

- Pour les items à réponse choisie, la clé de correction indique la bonne réponse, généralement par une lettre.
- Pour les items à réponse construite, la clé de correction fournit la bonne réponse ou encore un exemple de bonne réponse et/ou les caractéristiques que doit posséder la bonne réponse.

Exemple d'un item à réponse choisie :

Stimulus

Cent cinquante étudiants interrogés décrivent leurs activités après une journée de classe :
- 80 écoutent la radio
- 70 étudient leurs leçons
- 45 pratiquent un sport

Parmi tous ces étudiants, 25 écoutent la radio et étudient leurs leçons, 20 pratiquent un sport et étudient leurs leçons, 20 écoutent la radio et pratiquent un sport et enfin 5 pratiquent les trois activités.

Combien d'étudiants pratiquent une seule activité?

Choix de réponses

A) 10
B) 20
C) 40
D) 80

Clé de correction

D

2.1.2 Les règles de rédaction communes à tous les types d'items

Pour chacun des types d'items, il existe une série de règles à suivre et à respecter lors de la rédaction. Avant d'aborder les règles propres à chacun des types d'items, il convient de présenter celles qui sont communes à tous.

1) Tenir compte de la partie de la définition de domaine présentant les principes à respecter en évaluation, car certains d'entre eux influent directement sur la rédaction des items.

2) Rédiger les items de façon claire, dans une langue correcte, accessible aux élèves. Le vocabulaire doit être simple, autant que possible exempt de mots techniques.

3) S'assurer que le stimulus est complet : l'élève doit savoir exactement ce qu'on attend de lui. Il est préférable de présenter la question sous la forme interrogative plutôt que sous celle, énonciative, d'une phrase à compléter.

4) Dans le stimulus, éviter de fournir à l'élève des indices de réponses comme des énoncés empruntés textuellement aux manuels scolaires ou aux notes de classe, ou encore, des déterminants employés à la forme féminine ou plurielle, etc.

5) Dans le stimulus, éviter d'utiliser des expressions comme "selon toi", "d'après toi", "penses-tu que...", qui peuvent suggérer à l'élève qu'on lui demande son opinion.

6) Éviter de rédiger des phrases comprenant des négations ou des doubles négations.

7) Si on a recours à un extrait de document, en indiquer la source.

8) Expliquer à l'élève la façon d'indiquer ou de celles d'inscrire sa réponse.

9) Dans l'ensemble d'une épreuve, éviter qu'une réponse à une question fournisse la réponse à une autre question.

10) Présenter à l'élève des situations différentes de celles qu'il a pu examiner ou traiter en classe, ou de celles qu'il a pu trouver dans son manuel scolaire.

2.1.3 Les items à réponse choisie

2.1.3.1 Les items à choix multiple

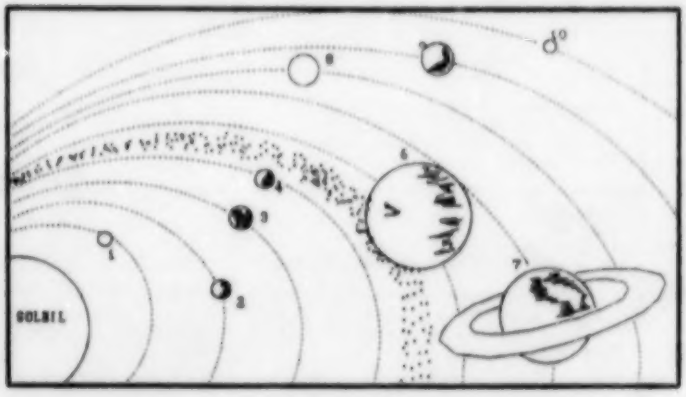
A) Définition

Dans l'item à choix multiple, on propose plusieurs réponses. L'élève choisit celle qu'il croit être la bonne. Les réponses inexactes sont appelées "leurres".

Exemple

Observe le schéma suivant de notre système solaire.

SYSTÈME SOLAIRE



Entre quelles planètes du système solaire la Terre se situe-t-elle?

A) 2 et 4
B) 3 et 5
C) 4 et 7
D) 5 et 8

Clé de correction

A

B) Avantages et limites des items à choix multiple

Tableau 6

Aspects examinés	Avantages	Limites
VALIDITÉ DE CONTENU	.Théoriquement, ils permettent d'évaluer les différents niveaux d'habiletés et ce, pour l'ensemble des programmes d'études. .Ces items assurent une bonne représentativité d'un programme d'études parce qu'il est possible d'en administrer un grand nombre en peu de temps.	.Ils permettent difficilement de mesurer les habiletés complexes.
FIDÉLITÉ	.L'effet du hasard est moindre que pour les items vrai ou faux. .Compte tenu de l'objectivité de la correction, ils sont plus fidèles que les items à réponse construite.	.Il est toujours possible que l'élève réponde au hasard.
CORRECTION	.Facile .Rapide .Objective .Elle peut se faire mécanographiquement.	
RÉDACTION	.Il est possible d'ajuster la difficulté d'un item en changeant les leurres. .Les leurres peuvent être rédigés de façon à permettre un diagnostic.	.Il est difficile de rédiger plusieurs leurres vraisemblables. .La rédaction prend beaucoup de temps. .La rédaction d'items qui mesurent des habiletés complexes demande beaucoup de savoir-faire.

C) Règles de rédaction des items à choix multiple

1) Tout en usant de concision, inclure le plus d'éléments possible dans le stimulus afin que les réponses soient aussi brèves que possible.

2) Rédiger les réponses proposées à la même forme grammaticale.

3) S'assurer d'établir un lien entre les réponses et le stimulus.

4) Proposer des réponses attrayantes et vraisemblables. On peut s'inspirer des erreurs commises le plus fréquemment par les élèves.

5) Éviter d'utiliser les expressions "aucune de ces réponses" ou "toutes ces réponses".

6) Séparer les réponses d'avec le stimulus. Placer les réponses selon un ordre logique (alphabétique ou numérique). En procédant de la sorte, la position de la bonne réponse variera de façon aléatoire, dans l'ensemble d'une épreuve.

7) Pour vérifier la connaissance d'une définition, placer le mot à définir dans le stimulus. Présenter les définitions dans les réponses.

8) Essayer de rédiger de trois à cinq réponses par choix de réponses. Il n'existe aucune règle qui spécifie que l'on doive toujours avoir le même nombre de réponses par choix dans une même épreuve. Il vaut mieux proposer trois réponses pertinentes et vraisemblables que cinq dont deux sont manifestement invraisemblables.

9) Éviter de présenter dans un choix deux réponses qui s'excluent mutuellement.

10) Proposer des réponses qui ont à peu près la même longueur.

11) S'assurer qu'il n'y a qu'une seule bonne réponse.

2.1.3.2 Les items vrai ou faux

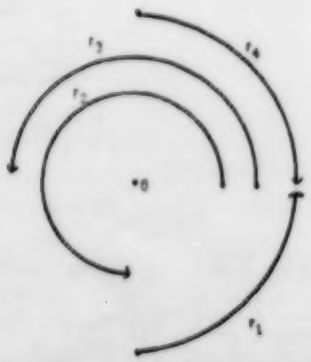
A) Définition

L'item vrai ou faux amène l'élève à choisir entre deux réponses contraires : vrai ou faux, oui ou non, en accord ou en désaccord, etc.

Exemple

Question 12

Vrai ou faux, parmi les rotations illustrées ci-dessous, la rotation r_4 est la seule rotation équivalente à r_3?



Réponse: _____

Clé de correction

Vrai

B) Avantages et limites des items vrai ou faux

Tableau 7

Aspects examinés	Avantages	Limites
VALIDITÉ DE CONTENU	.Ils permettent une bonne représentativité du programme d'études. En effet, pour une même période d'administration, un élève répond à trois fois plus d'items qu'avec des choix multiples. .Quand ils sont bien rédigés, ils permettent de mesurer plusieurs niveaux d'habiletés.	.Ailleurs qu'en mathématique et qu'en sciences, ils permettent difficilement de mesurer les habiletés complexes.
FIDÉLITÉ	.Compte tenu de l'objectivité de la correction, ces items sont plus fidèles que les items à réponse construite.	.Pour une fidélité égale à celle obtenue avec les items à choix multiple, il faut inclure deux fois plus d'items vrai ou faux. Ce type d'item permet moins bien de discriminer les élèves forts et les élèves faibles. . Plus le nombre d'items est restreint, plus l'effet du hasard est élevé.
CORRECTION	.Facile .Rapide .Objective .Elle peut se faire mécanographiquement.	
RÉDACTION		.Il est difficile de rédiger un item qui soit totalement vrai ou totalement faux. .Il est difficile de rédiger des items qui mesurent des habiletés complexes. .Ils ne permettent pas d'effectuer un diagnostic.

C) Règles de rédaction des items vrai ou faux

1) Énoncer des propositions compatibles avec l'alternative proposée (vrai ou faux, oui ou non, etc.).

2) Inclure une seule idée par énoncé à moins que l'on vérifie la relation entre deux éléments.

3) Écrire les énoncés à la forme affirmative.

4) Rédiger chacun des items de façon qu'il soit totalement vrai ou totalement faux.

5) Pour l'ensemble de l'épreuve, donner aux énoncés vrais et aux énoncés faux la même longueur afin d'éviter de fournir aux élèves des "patrons de réponses" qui leur serviraient d'indices.

R. Ebel (1972) propose des exemples de phrases types qui peuvent aider à rédiger des items vrai ou faux. Le tableau 8 présente quelques-unes de ces phrases types.

Tableau 8

PHRASES TYPES POUR LA RÉDACTION D'ITEMS VRAI OU FAUX

Si l'on veut vérifier l'habileté de l'élève à	Phrases types
généraliser :	. Tous les...
comparer :	. La différence entre... et... est... . ...nécessitent tous deux...
établir des relations :	. Le plus grand... . Le plus haut... . Le plus bas... . Tel... engendrera probablement... . L'augmentation de... tend à... . Combien de ... dépend de...
expliquer :	. La raison principale pour... . Le but de... . Un des facteurs qui affecte... . Étant donné que... . Même si...
prédire :	. On peut s'attendre à ce que... . De l'accroissement de... résultera...
vérifier une procédure :	. Pour trouver... on doit... . Pour faire... il faut... . La méthode pour... consiste en... . Une étape importante dans... est... . L'utilisation de... . L'étape essentielle pour... est...
évaluer :	. Il est préférable de... que de... . Le/la meilleur(e)... est... . Le maximum de... est... . Il est facile de démontrer que... . Il est difficile de... . Il est raisonnable de... . Il est nécessaire de... pour que... . L'inconvénient majeur de... est...

2.1.3.3 Les items de type appariement

A) Définition

L'item de type appariement présente d'abord une série d'énoncés qui constituent l'ensemble-stimuli : à chacun des éléments de ce premier ensemble, l'élève doit associer un élément faisant partie d'un deuxième ensemble, l'ensemble-réponses. Tous les éléments doivent référer à une même structure ou à un même ensemble de faits.

Exemple

À quel climat correspond chacun des regroupements de caractéristiques suivants? Pour répondre, inscris la lettre à côté de chaque numéro. Une même réponse ne peut être utilisée plus d'une fois.

CARACTÉRISTIQUES		CLIMATS
1.		
Latitudes	: 10° à 30°N	A) Continental
Moyenne de température en été	: 28°C	B) Équatorial
Moyenne de température en hiver	: 24°C	C) Méditerranéen
Formation végétale	: savane	D) Océanique
Espèces d'arbres ou de plantes	: baobab, teck	E) Tropical
Espèces d'animaux	: lion, gazelle, éléphant	
Activités particulières	: arbres fruitiers	
2.		
Latitudes	: 30° à 45°N	
Moyenne de température en été	: 26°C	
Moyenne de température en hiver	: 12°C	
Formation végétale	: garrigue, maquis	
Espèces d'arbres ou de plantes	: olivier, vigne, chêne-liège	
Espèces d'animaux	: animaux domestiques, lézard	
Activités particulières	: viticulture, culture des arbres fruitiers	
3.		
Latitudes	: 40° à 60°N/ 35° à 45°S	
Moyenne de température en été	: 21°C	
Moyenne de température en hiver	: 7°C	
Formation végétale	: forêt de feuillus	
Espèces d'arbres ou de plantes	: chêne, hêtre, tremble	
Espèces d'animaux	: animaux domestiques, lièvre, castor	
Activités particulières	: pêcheries	

Réponse

1. ( )
2. ( )
3. ( )

Clé de correction

1. E
2. C
3. D

B) Avantages et limites des items de type appariement

Tableau 9

Aspects examinés	Avantages	Limites
VALIDITÉ DE CONTENU	.Ils peuvent permettre de mesurer la compréhension des concepts ou des principes à la condition expresse de présenter à l'élève une situation nouvelle.	.Ils ne peuvent pas être utilisés pour mesurer les habiletés complexes. .Ils ne peuvent pas être utilisés quand le contenu notionnel est restreint. .Compte tenu du temps nécessaire pour répondre, ce type d'items, à lui seul, ne permet pas d'assurer la représentativité du programme.
FIDÉLITÉ	.Tout comme les items à choix multiple, les items de type appariement sont fidèles et ils permettent de bien discriminer les élèves forts et les élèves faibles.	
CORRECTION	.Facile .Rapide .Objective .Elle peut se faire mécanographiquement.	.Le nombre d'appariements est limité lorsque la correction est mécanographique.
RÉDACTION	.Ils sont appropriés dans les cas où l'utilisation du choix multiple obligerait à la répétition des mêmes choix de réponses.	.Il est difficile de faire en sorte que les stimuli soient de même nature. .Il est difficile de faire en sorte que chacune des réponses soit un choix vraisemblable pour chacun des stimuli.

C) Règles de rédaction des items de type appariement

1) Utiliser des chiffres pour identifier les stimuli et des lettres pour identifier les réponses.

2) Placer l'ensemble-stimuli dans une colonne, l'ensemble-réponses dans une autre. Donner un titre à chaque colonne.

3) Placer les choix de la colonne des stimuli selon un ordre logique (alphabétique ou numérique); faire la même chose pour les réponses.

4) Éviter la parité des appariements, c'est-à-dire un nombre de stimuli égal à celui des réponses. Généralement, il y a au moins trois stimuli. Il doit y avoir toujours plus de réponses que de stimuli. Il n'est pas recommandé de faire plus de cinq appariements par item.

5) Les contenus de l'ensemble-stimuli et de l'ensemble-réponses doivent être homogènes. Chaque réponse doit constituer un choix possible pour chaque stimulus.

6) Expliquer à l'élève comment se fait l'appariement.

7) Indiquer si une même réponse peut être utilisée plus d'une fois.

2.1.4 Les items à réponse construite

2.1.4.1 Les items à réponse courte

A) Définition

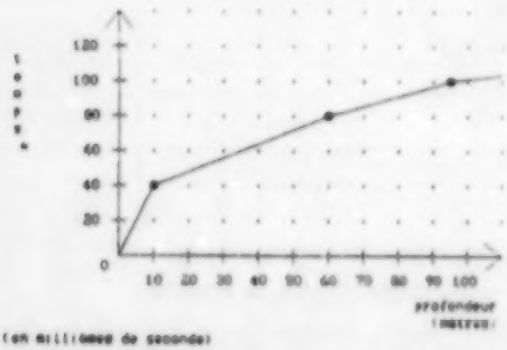
L'item à réponse courte est celui dont le stimulus précis conduit l'élève à fournir une réponse unique, brève et spécifique. Ce peut être un mot (mot de vocabulaire, nom de lieu, de personne, etc.), un chiffre, un symbole ou une courte phrase : une partie de définition, un énoncé.

L'élève doit se rappeler ou trouver la réponse, et l'écrire lui-même dans l'espace prévu à cette fin.

Exemple

Un géologue fait exploser une petite charge de dynamite à la surface du sol et enregistre, à l'aide d'un géophone, le temps que prennent les ondes de l'explosion pour traverser les différentes couches du sol, afin de connaître ainsi leur profondeur. Les enregistrements permettent de tracer le graphique suivant.

Première explosion



* (en millièmes de seconde)

Les changements de direction de la courbe du graphique indiquent des changements de vitesse des ondes de l'explosion lesquels se produisent aux changements de couches du sol.

1re couche traversée : terre
2e couche traversée : calcaire
3e couche traversée : granit

Détermine, par rapport à la surface du sol, la profondeur maximale de la couche de granit.

Réponse: _____

Clé de correction

95 mètres

B) Avantages et limites des items à réponse courte

Tableau 10

Aspects examinés	Avantages	Limites
VALIDITÉ DE CONTENU	.On obtient une meilleure validité qu'avec les items à réponse élaborée car on assure une meilleure représentativité du programme.	.Ils sont rarement appropriés pour mesurer les habiletés complexes. .Étant donné le temps requis pour y répondre, ils assurent une moins bonne représentativité du programme que les items à réponse choisie.
FIDÉLITÉ	.On obtient une fidélité supérieure à celle obtenue avec l'item à réponse élaborée, car la correction est à toutes fins pratiques objective puisque la réponse est généralement unique.	.La correction peut être biaisée par la perception de la personne qui corrige en regard de la calligraphie et de l'orthographe de l'élève de même que par sa connaissance de l'élève.
CORRECTION		.Une réponse peut être correcte sans toutefois correspondre à la réponse attendue. La mise à l'essai permet de réduire ou d'éliminer le problème en permettant de remplir la clé de correction à la lumière des réponses obtenues. .Il est nécessaire de transcrire les résultats d'élèves si l'on désire réaliser des analyses statistiques.
RÉDACTION	.La conception de ce type d'items est plus simple que celle des items à réponse choisie et également moins longue.	.On doit porter une attention particulière à la formulation du stimulus afin d'amener une réponse unique.

C) Règles de rédaction des items à réponse courte

1) Faire porter le stimulus sur un contenu très restreint et le rédiger de telle sorte qu'il n'existe qu'une seule bonne réponse et qu'elle s'exprime par un seul élément (mot, expression, etc.). Au besoin, indiquer toutes les réponses acceptables. Utiliser ce type d'items si la réponse peut être un mot ou une expression, un chiffre, un symbole. Ne pas choisir ce type d'items si une grande variété de réponses est possible, ou encore, si la longueur de la réponse dépasse une courte phrase.

2) Il n'est pas recommandé que la réponse comprenne plus de six éléments comme des symboles, des chiffres, des expressions, etc.

3) Si nécessaire, indiquer le degré de précision attendu (en mathématique et en sciences, surtout).

4) Au besoin, avertir l'élève qu'il sera pénalisé s'il introduit des éléments non pertinents dans sa réponse.

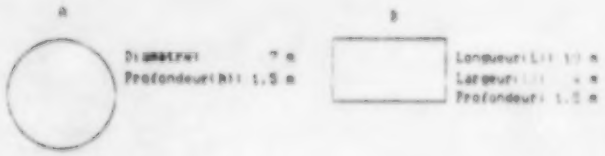
2.1.4.2 Les items à réponse élaborée

A) Définition

L'item à réponse élaborée est celui dont le stimulus présente un problème à résoudre et le cadre de référence nécessaire pour trouver la solution. L'élève doit résoudre le problème, structurer sa réponse (selon ses connaissances et le cadre de référence) et l'écrire avec ses propres mots. Chaque élève fournit une réponse différente, dont la pertinence, la justesse et la précision varient.

Exemple 1

Pour stimuler la vente de ses piscines, le propriétaire du magasin "Belleau" propose les piscines suivantes :



Le mélange des produits d'entretien (chlore et algicide) coûte 18,50 $ par mètre cube pour la saison estivale, combien économiseras-tu si tu achètes la piscine dont l'entretien est le plus économique?

Clé de correction

Démarche (exemple)

Volume de la piscine circulaire

$V_1 = \pi r^2 h$

$3,14 \times (3,5)^2 \times 1,5 = 57,7\ m^3$

Volume de la piscine rectangulaire

$V_2 = L \times l \times h$

$10 \times 4 \times 1,5 = 60\ m^3$

Différence des volumes

$V_1 - V_2 = 60 - 57,7 = 2,3\ m^3$

Différence de prix pour l'entretien

$2,3 \times 18,50 = 42,55$

Réponse

42,55 $

Exemple 2

TON PROJET

As-tu déjà vu un magicien exécuter des tours de magie?

Tes amis aimeraient connaître tes expériences.

Écris-leur un texte dans lequel tu décris un tour de magie.

Tu dois décrire les éléments suivants:

- L'endroit où le tour de magie a eu lieu.
- Le magicien (son habillement, ses gestes, sa façon de parler).
- Le tour de magie.
- Les objets utilisés par le magicien.

- Rédige d'abord un brouillon.
- Utilise ensuite la grille de révision et corrige ton texte.
- Retranscris ton texte au propre.
- Tu peux utiliser ton dictionnaire et ton cahier de règles de grammaire.

GRILLE DE RÉVISION

	OUI	NON
1. As-tu dit à quel endroit a eu lieu le tour de magie?	____	____
2. As-tu présenté le magicien?	____	____
3. As-tu parlé de son habillement?	____	____
4. As-tu parlé de ses mimiques?	____	____
5. As-tu parlé de sa façon de parler?	____	____
6. As-tu bien décrit tous les objets qui ont été utilisés?	____	____
7. As-tu décrit le tour de magie?	____	____
8. Tes phrases sont-elles bien structurées et bien ponctuées?	____	____
9. As-tu vérifié l'orthographe des mots?	____	____
10. As-tu appliqué les règles que tu as apprises?	____	____

Si tu as répondu NON à une ou plusieurs questions, tu dois faire les corrections appropriées pour améliorer ton texte.

B) Avantages et limites des items à réponse élaborée

Tableau 11

Aspects examinés	Avantages	Limites
VALIDITÉ DE CONTENU	.Ils permettent de mesurer les habiletés complexes. .Ils permettent d'évaluer la démarche de l'élève.	.Compte tenu de la longueur de la réponse et du temps nécessaire pour répondre, il est impossible d'inclure plusieurs items de ce type dans un même examen. Ce type d'items ne permet pas d'assurer la représentativité de l'épreuve en regard du programme.
FIDÉLITÉ	.Les chances de réussir un item au hasard sont presque nulles.	.La fidélité est très faible car les scores sont affectés par plusieurs facteurs tels que la calligraphie, l'orthographe, le style de l'élève. .La correction peut varier d'un correcteur ou d'une correctrice à l'autre. L'effet de "halo" intervient : inconsciemment, la personne qui corrige cote mieux les élèves qu'elle considère les meilleurs, et inversement; il arrive qu'elle cote différemment selon le moment où elle corrige les items.
CORRECTION		.L'objectivité de la personne qui corrige est très difficile à assurer. .Le temps de correction est très long. .Il faut transcrire les résultats des élèves si on désire réaliser une analyse statistique des résultats.
RÉDACTION	.La conception de ce type d'items est plus simple que celle des items à réponse choisie et moins longue.	.Trop souvent, la formulation n'est pas claire et les éléments de réponse sont mal définis. .La clé de correction se doit d'être très explicite. Il n'est pas facile de prévoir toutes les réponses. .L'élaboration de grilles d'observation est complexe.

C) Règles de rédaction des items à réponse élaborée

1) Structurer le stimulus de manière à orienter la réponse vers ce que l'on souhaite mesurer. Préciser le cadre de la réponse; bien délimiter et circonscrire le contenu. Le stimulus doit comporter un contexte ou une mise en situation présentant à l'élève toutes les informations dont il a besoin pour effectuer la tâche.

2) Présenter le stimulus par écrit. Toutefois éviter de rédiger un texte trop long renfermant des informations superflues quant à la tâche demandée.

3) Indiquer à l'élève la longueur de la réponse attendue, les critères de correction, les points alloués, le temps prévu pour répondre.

4) Ne pas privilégier dans le stimulus les formulations du type "quoi-qui-quand-où", qui conviennent mieux aux items à réponse courte ou à choix multiple.

5) Préparer soigneusement le contenu de la clé de correction : la grille d'observation, s'il y a lieu (voir 2.1.4.3), et le nombre de points alloués. Il s'agit de prévoir le plus précisément possible les caractéristiques de la réponse attendue.

6) Ne pas proposer aux élèves de choix entre plusieurs questions à réponse élaborée. En effet, il n'est pas certain que les questions soient d'égale difficulté. Les élèves auront du mal à choisir; les bons élèves risquent d'être pénalisés car ils choisiront sans doute la question la plus difficile, celle qui offre le plus grand défi. La comparaison des résultats en sera rendue presque impossible. Il est préférable que tous les élèves répondent à la même question.

D) Méthodes de correction

Compte tenu des problèmes de fidélité propres à ce type d'items, il est nécessaire de prévoir des mesures pour tenter d'uniformiser le plus possible la correction, surtout s'il s'agit d'une épreuve servant à des fins sommatives. Deux méthodes sont suggérées: la méthode analytique et la méthode globale.

1. La méthode analytique

Il est suggéré d'avoir recours à la méthode analytique lorsque le nombre de copies à corriger n'est pas trop élevé. L'utilisation de cette méthode, surtout si la correction est réalisée par les mêmes personnes, facilite la comparaison des résultats. Toutefois, le temps de correction est considérablement long.

Voici en quoi la méthode consiste :

1. Rédiger la réponse idéale en déterminant les aspects à contrôler, soit les éléments de réponse attendus. Construire, s'il y a lieu, une grille d'observation.

2. Lire quelques réponses fournies par les élèves et ajuster au besoin la clé de correction.

3. Lire d'abord toutes les copies d'élèves, puis corriger un item ou un critère à la fois.

4. Si possible, ne pas corriger les copies de son propre groupe d'élèves. Échanger les copies avec un autre correcteur ou une autre correctrice afin de diminuer l'effet de "halo", c'est-à-dire l'influence de la connaissance d'un élève sur la personne qui corrige et sur le jugement qu'elle portera.

5. Corriger s'il y a lieu l'orthographe ou la syntaxe, indépendamment du contenu de la réponse attendue, sauf en langues lorsque ce contenu est précisément de nature orthographique et syntaxique.

2. La méthode globale

Il est suggéré d'utiliser la méthode globale lorsque le nombre de copies est considérable. Le temps de correction n'est pas trop long. Voici en quoi cette méthode consiste :

1. Les copies ne sont pas corrigées à l'aide d'une grille détaillée ou d'une grille d'observation, mais plutôt globalement.

   Déterminer au départ les principaux critères de correction ou les éléments de réponse attendus. Tous les critères de correction sont considérés en même temps pour juger de la réponse.

2. Selon un premier jugement, lire toutes les copies, puis les partager en quatre ou cinq groupes.

3. Relire une deuxième fois les cas frontières pour les classer.

4. Faire une dernière lecture afin d'attribuer des notes selon un continuum de 0 à 100 % ou selon une échelle descriptive.

2.1.4.3 Les grilles d'observation

A) Définition

Une grille d'observation est un instrument de mesure qui permet de constater les particularités d'une action, d'un produit ou d'un processus en fournissant une liste d'éléments observables ainsi qu'une façon d'enregistrer les observations. Le terme grille d'observation englobe les différents types de grilles : grilles de correction pour les productions écrites, grilles d'appréciation pour les productions orales ou pour d'autres types de produits tels qu'un dessin ou un objet construit, grilles d'observation en classe pour comprendre le fonctionnement cognitif d'un élève face à une tâche.

B) Avantages et limites des grilles d'observation

1. Les avantages

Toutes les grilles d'observation présentent un grand intérêt pour l'évaluation formative. En effet, elles permettent d'identifier les éléments d'apprentissage qui nécessitent des correctifs. Dans le cas des productions écrites et des productions orales, on utilise aussi les grilles d'observation pour l'évaluation sommative.

2. Les limites

La construction et l'utilisation de grilles d'observation demandent beaucoup de temps. Les résultats des recherches sur la validité et la fidélité de ces instruments sont décevants. Il faut bien admettre également que les principes sur lesquels repose la construction des grilles d'observation sont peu nombreux. Cependant, leur pertinence ne peut être remise en cause dans un contexte formatif et même dans un contexte sommatif pour les réponses construites. Bien qu'imparfaits, ces outils donnent une qualité d'informations que l'on n'obtient pas autrement.

C) Règles de construction des grilles d'observation

Les grilles d'observation sont composées de critères et d'une échelle d'appréciation.

1. Sélectionner les critères

Il s'agit d'abord de demander à des spécialistes de la matière de dresser une liste de comportements observables chez un élève, comportements qui expriment l'action, le produit et le processus que l'on veut observer au moyen d'une grille. Ces listes sont ensuite mises en commun et, par regroupements et recoupements, on effectue une première sélection de critères. Celle-ci est généralement trop abondante, car on y retrouve de nombreux critères qui recouvrent les mêmes comportements sous des appellations différentes. Il faut donc procéder à une nouvelle analyse de contenu et à de nouveaux élagages avant de parvenir à la sélection définitive.

Il n'y a pas de formule pour déterminer le nombre de critères à retenir. La seule règle importante à respecter est la suivante : les critères retenus doivent décrire de façon exhaustive ce que l'on veut mesurer et être indépendants l'un de l'autre.

Toutefois, il ne faut pas perdre de vue que cet instrument doit être d'une utilisation pratique tout en étant le plus objectif possible.

On peut également avoir recours à des analyses statistiques pour systématiser la sélection des critères. Des spécialistes sont d'abord invités à décrire de la façon la plus précise possible une performance d'élève. Les descriptions ainsi obtenues sont alors soumises à une analyse de contenu, de laquelle on tire une liste de critères. La deuxième étape consiste à expérimenter ces critères auprès de la population cible et à soumettre les résultats à des analyses statistiques qui permettent de retenir une liste de critères indépendants.

2. Pondérer les critères

On peut soit considérer les critères comme des unités équivalentes, soit leur accorder une importance différente. Dans ce dernier cas, les spécialistes devront établir une procédure pour établir la pondération de chacun des critères puis valider ces pondérations auprès d'un nombre suffisant de juges. S'il existe des grilles d'observation pour lesquelles les pondérations ont déjà été validées, la procédure de validation perd de son importance : on peut utiliser ces pondérations comme base de travail et les adapter au besoin. Par exemple, dans tous les documents Développement de l'habileté à écrire. Exigences minimales au primaire, on trouve un tableau présentant la pondération de chacun des critères de correction des productions écrites pour toutes les classes de cet ordre d'enseignement.

3. Définir les critères

La description des critères a pour but de favoriser leur compréhension univoque par tous les utilisateurs de la grille d'observation. Pour préciser un critère, on rédige une courte description. Les verbes d'action sont conjugués au présent et à la forme affirmative.

Exemple

Extrait d'une grille d'appréciation d'une production orale à caractère incitatif, primaire, 2e cycle.

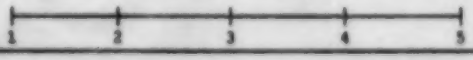
CRITÈRES
1. L'élève donne des informations en rapport avec le sujet.
2. L'élève utilise des arguments convaincants.
3. Les phrases sont bien construites.
4. Etc.

4. Définir l'échelle d'appréciation

L'échelle d'appréciation est la partie de la grille d'observation qui se présente sous la forme d'un continuum et qui permet d'indiquer la qualité ou la quantité des comportements décrits. Il existe différentes sortes d'échelles d'appréciation : l'échelle numérique, l'échelle graphique et l'échelle descriptive. Cette dernière est de loin la plus fiable parce qu'elle encadre la tâche de la personne qui évalue et qu'elle précise la description du critère.

Exemples

CRITÈRE	ÉCHELLE GRAPHIQUE
L'élève utilise un vocabulaire précis et riche.	

CRITÈRE	ÉCHELLE NUMÉRIQUE
L'élève utilise un vocabulaire précis et riche.	

CRITÈRE	ÉCHELLE DESCRIPTIVE				
	A	B	C	D	E
L'élève utilise un vocabulaire précis et riche.	Toujours précis et riche 10	Assez précis et riche 8	Parfois imprécis et ordinaire 6	Souvent imprécis et pauvre 4	Toujours imprécis et pauvre 0

Pour construire une échelle descriptive, on définit d'abord les points extrêmes. On formule ensuite la description des points plus au centre en s'efforçant de distancer les points le plus également possible. Quand on décrit une qualité, les descriptions sont précises, concises et rédigées avec la même forme grammaticale. Si on utilise des phrases, elles sont au présent et à la forme affirmative.

5. Déterminer le nombre d'échelons

L'échelle à cinq échelons est sûre et elle permet l'utilisation d'une cote moyenne. C'est un avantage à considérer, sans toutefois oublier que les utilisateurs ont tendance à choisir le troisième échelon. L'échelle à quatre échelons est d'une application plus rapide alors que celle à six échelons offre une plus grande gamme de nuances; par contre, cette dernière est plus difficile à construire et à utiliser. Les recherches ont démontré que les échelles à quatre et celles à six échelons se différencient peu du point de vue de la fidélité et de la sensibilité. Souvent, il suffit de trois points de repère pour constituer une échelle d'utilisation facile et efficace en classe.

Comme on peut le constater, le nombre d'échelons sera choisi en fonction du but recherché. L'essentiel est d'opter pour l'échelle descriptive et de définir de façon précise chacun des échelons.

6. Assembler la grille d'observation

Bien que l'assemblage de la grille ne revête pas la même importance que les aspects précédents, il peut s'avérer utile d'énumérer un certain nombre d'éléments d'ordre technique qui facilitent la tâche de l'utilisateur et qui, dans certains cas, contribuent à augmenter la fidélité de la grille d'observation. Ces éléments sont regroupés ici en deux catégories : ceux qui concernent la grille d'observation elle-même et ceux qui ont trait au guide de l'enseignant ou de l'enseignante.

A) <u>La grille d'observation</u>

- Soigner la présentation visuelle tout en essayant de regrouper les critères à observer sur une même feuille.
- Présenter, quand cela est nécessaire, les critères à observer dans l'ordre d'apparition des comportements à observer.
- Prévoir une méthode facile et rapide de consigner le résultat des observations (X, ✓, etc.).
- Prévoir un espace, s'il y a lieu, pour consigner le résultat global.
- Prévoir un espace pour inscrire le nom de l'élève. Dans certains cas, il peut être intéressant de prévoir un espace pour la date et un pour le nom de la personne qui procède à l'observation.

B) <u>Le guide de l'enseignant ou de l'enseignante</u>

Il est important de préparer un guide pour l'enseignant ou l'enseignante qui utilisera la grille d'observation. Ce guide réunit, d'une part, les directives d'administration et, d'autre part, des précisions relatives aux critères -- une partie importante du guide -- et des explications complémentaires sur l'échelle descriptive.

Chaque fois que cela est pertinent, on peut ajouter des commentaires ou des exemples pour préciser la description des échelons. On décrit par ailleurs dans ce guide la méthode de notation ou d'observation à utiliser. L'utilisation d'un tel guide contribue à augmenter la fidélité des observations.

7. Expérimenter la grille d'observation

L'expérimentation auprès de la clientèle cible permet de déceler les vices de construction de la grille et les erreurs faites par les utilisateurs de la grille. Ce genre d'opération nécessite l'organisation de séances de correction collective ou d'observation simulée. Ces rencontres permettent de déceler les faiblesses de l'instrument à partir des résultats ou des commentaires des utilisateurs. Il y a lieu aussi, lors de ces rencontres, de prévoir du temps pour questionner ces personnes sur leur façon de comprendre et d'utiliser la grille. Cela est particulièrement utile lorsqu'il s'agit d'une nouvelle grille d'observation ou pour des domaines dans lesquels l'observation n'est pas pratique courante.

Voici des exemples des vices de construction le plus fréquemment rencontrés:

- la formulation des critères ou la description des échelons sont ambiguës;
- les catégories de l'échelle ne sont pas exhaustives;
- les catégories de l'échelle ne sont pas mutuellement exclusives.

Quant aux erreurs faites par les utilisateurs, les plus fréquentes sont :

- interpréter les critères;
- ajouter un échelon à l'échelle;
- accorder une note sans utiliser l'échelle.

Si une expérimentation à grande échelle n'est pas réalisable, on peut la remplacer par un essai auprès de quelques groupes d'élèves seulement. A condition d'être systématique, cette expérience permettra de déceler les principaux problèmes. Idéalement, la grille d'observation est présentée aux utilisateurs, et des séances de correction collective ou d'observation simulée sont organisées régulièrement. Ces séances de relance permettent de mettre à jour la grille et d'assurer une compréhension plus uniforme.

12 RÉVISER ET CORRIGER LES ITEMS

Avant d'assembler l'épreuve, il importe de procéder à la validation de contenu et à la révision docimologique et à la correction des items.

Comme nous l'avons vu précédemment, les règles de rédaction d'items sont nombreuses. Il arrive souvent qu'une ou plusieurs règles échappent au rédacteur ou à la rédactrice; c'est pourquoi il est important de réviser les items. Les grilles présentées dans les pages suivantes peuvent servir de guide ou d'aide-mémoire au moment de la révision. Elles présentent l'essentiel des règles de rédaction.

Le mieux serait de demander à des collègues de lire les items afin de déceler les erreurs ou les ambiguïtés qui ont pu échapper à la personne qui rédige. Si cela est impossible parce que celle-ci est le seul ou la seule spécialiste de la matière dans l'école, on devra laisser s'écouler un certain temps entre l'étape de la rédaction et celle de la révision. Cela permettra au rédacteur ou à la rédactrice de devenir critique face aux items. On sélectionnera ensuite parmi les items élaborés ceux qui feront partie de l'épreuve.

GRILLE DE RÉVISION DES ITEMS				
Critères valables pour tous les types d'items	Oui	Non	Incertain	Ne s'applique pas
1. L'item est-il congruent avec sa dimension : mesure-t-il bien les contenus et les habiletés visés?				
2. Le type d'items retenu est-il approprié pour mesurer la dimension?				
3. L'item est-il clairement formulé?				
4. Le niveau de langue utilisé correspond-il au niveau de langue des élèves?				
5. Le stimulus est-il concis?				
6. A-t-on évité de fournir des indices de bonne réponse (adverbes de temps, nombre ou genre des déterminants)?				
7. Les textes et documents sont-ils adaptés aux élèves?				
8. La source des documents cités est-elle indiquée?				
9. La qualité des documents ou des dessins est-elle impeccable?				
10. A-t-on évité d'utiliser des extraits de manuels scolaires?				
11. L'item est-il exempt de stéréotypes sexistes, culturels ou autres?				

Note : Dans les cas où la réponse à ces questions est Non ou Incertain, il y a lieu de corriger l'item ou les items concernés.

GRILLE DE RÉVISION DES ITEMS				
Items à choix multiple	Oui	Non	Incertain	Ne s'applique pas
1. Les leurres sont-ils tous vraisemblables?				
2. Les réponses présentées sont-elles indépendantes l'une de l'autre?				
3. Les réponses de l'item sont-elles rédigées à la même forme grammaticale?				
4. Les réponses de l'item sont-elles de même longueur et de même complexité?				
5. Les réponses de l'item sont-elles placées dans un ordre logique (alphabétique ou numérique)?				
6. A-t-on évité l'utilisation des expressions «aucune de ces réponses» et «toutes ces réponses»?				

Note : Dans les cas où la réponse à ces questions est Non ou Incertain, il y a lieu de corriger l'item ou les items concernés.

GRILLE DE RÉVISION DES ITEMS				
Items de type appariement	Oui	Non	Incertain	Ne s'applique pas
1. Les contenus des stimuli et des réponses sont-ils homogènes?				
2. Est-ce que chacune des réponses constitue un choix vraisemblable pour chacun des stimuli?				
3. Est-ce que les stimuli et les réponses sont présentés dans un ordre logique (alphabétique ou numérique)?				
4. Y a-t-il plus de réponses que de stimuli?				
5. Le nombre de stimuli est-il inférieur à 5?				
6. A-t-on indiqué si une même réponse peut être utilisée plus d'une fois?				
7. Est-ce que les colonnes des stimuli et des réponses ont été identifiées?				
8. L'élève sait-il de quelle façon répondre?				

Note : Dans les cas où la réponse à ces questions est Non ou Incertain, il y a lieu de corriger l'item ou les items concernés.

GRILLE DE RÉVISION DES ITEMS				
Items à réponse courte et items à réponse élaborée	Oui	Non	Incertain	Ne s'applique pas
1. Le contenu du stimulus est-il très restreint, amenant, dans le cas des items à réponse courte, à une réponse unique?				
2. Le degré de précision attendu (si nécessaire) est-il mentionné à l'élève?				
3. Dans le cas des items à réponse élaborée, le stimulus comporte-t-il un contexte ou une mise en situation présentant à l'élève toutes les informations dont il a besoin pour répondre?				
4. La longueur de la réponse attendue, les critères de correction et les points à allouer sont-ils précisés?				
5. Dans la clé de correction, toutes les réponses acceptables sont-elles indiquées?				
6. A-t-on indiqué, dans la clé de correction, les conditions de réussite (nombre d'éléments de réponse attendus pour réussir l'item)?				

Note : Dans les cas où la réponse à ces questions est Non ou Incertain, il y a lieu de corriger l'item ou les items concernés.

GRILLE DE RÉVISION DES ITEMS				
Critères pour l'ensemble de l'épreuve	Oui	Non	Incertain	Ne s'applique pas
1. Le but de l'épreuve est-il clairement établi?				
2. L'épreuve couvre-t-elle l'ensemble des dimensions retenues dans le tableau de la définition de domaine?				
3. Les pondérations indiquées dans le tableau des dimensions ont-elles été respectées?				
4. Les principes à respecter, énoncés dans la définition de domaine, l'ont-ils été?				
5. Le nombre de points à accorder à chaque item est-il indiqué?				
6. La réponse est-elle clairement indiquée pour chacun des items?				
La réponse indiquée est-elle correcte et la seule correcte?				
7. L'ordre des items est-il pertinent?				
8. La longueur de l'épreuve est-elle appropriée?				
9. Les directives pour l'élève sont-elles claires, complètes, précises et bien en évidence?				
10. A-t-on évité de couper un item en le répartissant sur deux pages?				
11. L'épreuve est-elle exempte d'erreurs grammaticales, orthographiques et typographiques?				
12. Les caractères d'imprimerie sont-ils lisibles?				

Note : Dans les cas où la réponse donnée aux questions est Non ou Incertain, il y a lieu de corriger le ou les critères concernés.

CHAPITRE 3

ASSEMBLER ET ADMINISTRER L'ÉPREUVE

3.1 ASSEMBLER L'ÉPREUVE

Au moment d'assembler l'épreuve, il faut disposer des documents suivants :

- le cahier de l'élève;
- le cahier ou la feuille de réponses;
- les directives d'administration;
- la clé de correction;
- le guide de l'enseignant ou de l'enseignante (s'il y a lieu).

3.1.1 Le cahier de l'élève

Ce cahier contient tous les items auxquels l'élève doit répondre. Les items sont précédés des directives pour l'élève.

3.1.1.1 L'ordre des items

Il convient d'ordonner les items selon leurs diverses caractéristiques : leur difficulté, leur type, les thèmes qu'ils abordent.

La plupart des spécialistes en mesure et évaluation s'entendent pour dire qu'il importe de placer les items dans un ordre croissant de difficulté. Cela permet à la majorité des élèves de répondre rapidement et correctement aux premiers items sans risquer d'être découragés dès le départ par des items trop difficiles; ils se libèrent ainsi peu à peu de l'anxiété inhérente au fait de passer un test. Leur temps et leur effort s'en trouvent répartis plus efficacement tout le long de l'épreuve.

Cependant, il n'est pas souhaitable que l'élève ait à changer trop souvent sa façon de répondre, ce qui le retarderait et pourrait lui occasionner des erreurs. Il serait donc à propos de regrouper les items de même type: ceux à réponse choisie, les plus faciles, au début (en plaçant ensemble respectivement les items à choix multiple, les items vrai ou faux, ceux de type appariement), et les items à réponse construite à la fin.

Une remarque semblable s'applique aux thèmes : un mélange continuel ne ferait qu'embrouiller l'élève inutilement. Il est préférable de réunir tout ce qui concerne un même thème (par exemple : les nombres, la géométrie, la mesure), en prenant soin de disposer les items, à l'intérieur du thème, par ordre croissant de difficulté.

Enfin, dans une épreuve qui porte sur un document, par exemple sur un texte en français, surtout s'il est long, il peut être impératif de suivre l'ordre du document afin d'éviter à l'élève un va-et-vient superflu.

Dans tous les cas où l'ordre croissant de difficulté n'est pas respecté du début à la fin de l'épreuve, il importe de recommander aux élèves de ne pas s'attarder sur un item qu'ils jugent difficile, mais de passer plutôt à d'autres parties de l'épreuve sans manquer toutefois, si le temps le permet, de revenir sur les items auxquels ils n'ont pas répondu.

3.1.1.2 Les directives pour l'élève

Les directives pour l'élève, préférablement placées en page couverture du cahier, doivent être explicites.

Il faut indiquer clairement à l'élève comment répondre aux divers types d'items. Un exemple peut être fourni, particulièrement si l'épreuve s'adresse aux élèves du primaire.

Il faut informer l'élève de la pondération accordée à chaque item. Si les items ont une valeur égale, il suffira de la mentionner dans les directives pour l'élève; sinon, il est mieux d'inscrire la valeur de chaque item dans le cahier de l'élève, dans la marge, vis-à-vis le numéro de l'item.

Il faut toujours indiquer la durée prévue pour l'épreuve. D'autres directives peuvent également apparaître, notamment le matériel autorisé, le matériel non autorisé, etc.

3.1.1.3 Le cahier ou la feuille de réponses

Le fait de fournir un cahier ou une feuille de réponses à l'élève réduit considérablement le temps de correction. En effet, plutôt que d'avoir à tourner toutes les pages du cahier de l'élève, et ce, pour chaque élève, l'enseignant ou l'enseignante ne manipule que la ou les feuilles de réponses. C'est plus simple, plus rapide, sans compter que cela réduit le risque d'erreurs, comme l'omission de pages ou de numéros. Si l'épreuve est composée uniquement d'items à réponse choisie et qu'un lecteur optique est disponible, on peut utiliser des feuilles mécanographiques.

Au primaire, particulièrement au 1er cycle, il est préférable de ne pas utiliser de cahier ou de feuille de réponses afin de ne pas compliquer la tâche de l'élève.

3.1.2 Le guide de l'enseignant ou de l'enseignante

On peut présenter séparément les directives d'administration et la clé de correction de l'épreuve, ou les réunir dans un guide de l'enseignant ou de l'enseignante qui devrait aussi, idéalement, comprendre le tableau des dimensions de la définition de domaine indiquant la répartition des items.

Les directives d'administration comprennent :

. la durée de l'épreuve;
. le matériel requis;
. les conditions d'administration.

La clé de correction comprend :

. les réponses attendues;
. le nombre de points par item;
. les critères permettant d'attribuer les points, s'il y a lieu.

3.2 ADMINISTRER L'ÉPREUVE

3.2.1 Mettre à l'essai et corriger l'épreuve

La mise à l'essai d'une épreuve est une étape très importante puisqu'elle permet de déceler les problèmes que peuvent poser certains items. Cependant, elle n'est pas toujours possible, par exemple dans le cas où les enseignants travaillent seuls.

En effet, il faut à cette fin trouver un ou deux groupes d'élèves autres que ses propres groupes. Cet échantillon doit ressembler le plus possible à la population d'élèves à laquelle l'épreuve sera administrée ultérieurement. À la lumière des résultats obtenus, des commentaires des enseignants qui ont participé et des commentaires des élèves faisant partie de l'échantillon, on relève les items défectueux afin d'apporter les corrections appropriées. Par exemple, si tous les élèves ont eu besoin d'informations supplémentaires pour répondre à une question, on peut penser que la formulation de cette question est à retravailler.

3.2.2 Expérimenter l'épreuve

Une fois les corrections apportées, après l'essai, on administre l'épreuve à un plus large échantillon d'élèves afin de pouvoir effectuer des analyses statistiques. L'interprétation des résultats est importante, car elle fait ressortir les items défectueux. Par exemple, il est possible que l'on doive retirer certains items; dans ce cas, les points alloués pour ces items ne doivent pas être inclus dans le score brut de l'élève.

Les personnes voulant obtenir de plus amples renseignements sur la procédure à suivre pour l'analyse des résultats sont invitées à se référer au document Exploitation des données de la mesure du ministère de l'Éducation.

BIBLIOGRAPHIE

AUGER, Réjean et Michel FRECHETTE. "La définition de domaine, une étape essentielle dans l'élaboration d'un instrument de mesure", *Mesure et évaluation en éducation*, vol. 10, n° 4, hiver 1988.

BERK, Ronald A. *Criterion-referenced Measurement : The State of the Art*, The Johns Hopkins University Press, 1989, 233 p.

BERNIER, Jean-Jacques. *Théorie des tests, principes et techniques de base* Chicoutimi, Gaëtan Morin Éditeur, 1984, 275 p.

BLOOM, Benjamin S., HASHINGS J. Thomas et George F. MADAUS. *Evaluation to Improve Learning*, New York, McGraw-Hill Book, 1981.

BLOOM, Benjamin S. et autres. *Taxonomie des objectifs pédagogiques*, traduit par Marcel Lavallée, Montréal, Éducation nouvelle, 1969, 233 p.

CARDINET, Jean. *Pour apprécier le travail des élèves*, Bruxelles, De Boeck-Wesmael, s.a., 1986, 131 p.

CAREY, Lou M. *Measuring and Evaluating School Learning*, Allyn and Bacon, Inc., 1988, 456 p.

CUNNINGHAM, George K. *Educational and Psychological Measurement*, New York, Macmillan Publishing Co., 1986, 407 p.

DE BAL R., DE LANDSHEERE G. et J.P. BECKERS. *Construire des échelles d'évaluation descriptives*, Ministère de l'Éducation nationale et de la Culture franc[illegible]976, p. 69 à 107.

DE KETELE, Jean-Marie. *L'évaluation : approche descriptive ou prescriptive*, Bruxelles, De Boeck-Wesmael, s.a., 1986, 288 p.

DE LANDSHEERE, G. et V. DE LANDSHEERE. *Définir les objectifs de l'éducation*, 2e édition, Paris, P.U.F., 1976, 293 p.

DE LANDSHEERE. *Dictionnaire de l'évaluation et de la recherche en éducation*, Paris, P.U.F., 1979, 338 p.

D'HAINAULT, Louis. *Des fins aux objectifs de l'éducation*, Éducation 2000, 2e édition, Paris, Fernand Nathan, 1980.

DIENEY, Henry. *Classroom Evaluation for Teachers Issues and Innovations in Education*, Dubaque, Iowa, W.M.C. Brown Co. Publ., Dubaque, Iowa, 1971, 119 p.

EBEL, Robert L. et David A. FRISBIE. *Essentials of Educational Measurement*, 4th Ed., Prentice-Hall, Inc., 1986, 360 p.

HOGE, Robert D. "The Validity of Direct Observation Measure of Pupil Classroom Behavior", *Review of Educational Research*, hiver 1985, vol. 55, n° 4, p. 469-483.

HOPKINS, Charles D. et Richard L. ANTES. *Classroom Measurement and Evaluation*, F.E. Peacock Publishers, Inc., 1978, p. 465.

GOUVERNEMENT DU QUÉBEC. Ministère de l'Éducation. Direction de la coordination des réseaux. *Perfectionnement en évaluation pédagogique. L'exploitation des données de la mesure*, 1988.

GOUVERNEMENT DU QUÉBEC. Ministère de l'Éducation. Direction de la formation générale. *Document d'information. Définition du domaine. Histoire du Québec et du Canada*, version révisée, décembre 1986, 22 p.

GOUVERNEMENT DU QUÉBEC. Ministère de l'Éducation. Direction de la formation générale. *Document d'information. Définition du domaine. Mathématique, 1er cycle du primaire*, version automne 1988, 12 p.

GOUVERNEMENT DU QUÉBEC. Ministère de l'Éducation. Direction de la formation générale. *Document d'information. Définition du domaine. Mathématique, 1er cycle du secondaire*, version automne 1987, 8 p.

GOUVERNEMENT DU QUÉBEC. Ministère de l'Éducation. Direction de la formation générale. *Développement de l'habileté à écrire. Exigences minimales. Primaire, 1re à 6e année. 1986*.

GOUVERNEMENT DU QUÉBEC. Ministère de l'Éducation. Direction générale de l'évaluation. *Éléments de docimologie. fascicule 1 : L'évaluation pédagogique, une démarche*, 1985, 29 p.

GOUVERNEMENT DU QUÉBEC. Ministère de l'Éducation. Direction générale de l'évaluation. *Éléments de docimologie. fascicule 2 : Lexique*, 1985, 23 p.

GOUVERNEMENT DU QUÉBEC. Ministère de l'Éducation. Direction générale de l'évaluation. *Éléments de docimologie, fascicule 3 : L'évaluation formative*, 1985, 118 p.

GOUVERNEMENT DU QUÉBEC. Ministère de l'Éducation. Direction générale de l'évaluation. *Éléments de docimologie, fascicule 4 : L'évaluation sommative*, 1987, 92 p.

GOUVERNEMENT DU QUÉBEC. Ministère de l'Éducation. Direction générale de l'évaluation. *Guide d'évaluation en classe. Évaluation formative. Introduction générale. Primaire, secondaire formation générale, secondaire formation professionnelle*, 1985, 29 p.

GOUVERNEMENT DU QUÉBEC. Ministère de l'Éducation. Direction générale de l'évaluation. *Politique générale d'évaluation pédagogique*, Québec, 1981, 23 p.

GOUVERNEMENT DU QUÉBEC. Ministère de l'Éducation. Direction générale de l'évaluation. *Répertoire des publications de la Direction du développement de l'évaluation (1980-85)*, 1986, 14 p.

LEGENDRE, Renald. *Dictionnaire actuel de l'éducation*, Paris-Montréal, Larousse, 1988, 680 p.

MEHRENS, William A. et Irvin J. LEHMANN, *Measurement and Evaluation in Education and Psychology*, Third Edition, 1984, 658 p.

MESSICK, S. "Constructs and Their Vicissitudes in Educational and Psychological Measurement", *Psychological Bulletin*, 1981, p. 575-588.

MORISSETTE, Dominique. *Les examens de rendement scolaire. Comment les préparer et comment les administrer*, Québec, Les Presses de l'Université Laval, 1979. 390 p.

NITKO, Antony J. *Educational Tests and Measurement an Introduction*, New York, Harcourt Brace Jovanovich, Inc., 1983, 674 p.

OWEN, Steven V. et Robin D. FROMAN. "What's Wrong with Three Multiple Choice Items?" *Educational and Psychological Measurement*, vol. 47, 1987, p. 513 à 524.

ROID, Gale et Thomas, HALADYNA. "The Emergence of an Item-writing Technology", *Review of Educational Research*, été 1980, vol. 50, n° 2, p. 293-314.

SCALLON, Gérard. *L'évaluation formative des apprentissages, L'instrumentation*, Québec, Les Presses de l'Université Laval, 1988, 263 p.

SCALLON, Gérard. *L'évaluation formative des apprentissages*. *La réflexion*. Québec, Les Presses de l'Université Laval, 1988, 171 p.

SHROCK, Sharon, MANSUKKANI, Ramesh H., COSCARELLI, William et Sam, PALMER, [illegible]. "An Overview of Criterion-referenced Test Development", *Performance and Instruction*, vol. 25, n° 6, août 1986, p. 3-7.

STIÉVENART, Michel et Yvan TOURNEUR. «Modèles de production de questions à partir d'objectifs opérationnels. Évaluation au cours des dernières décennies», *Mesure et évaluation*, vol. 11, n° 3, automne 1988, p. 5 à 32.

THORNDIKE, Robert L. et Elizabeth P. HAGEN. *Measurement and Evaluation in Psychology and Education*, Fourth Edition, John Wiley & Sons, 1977, 693 p.

TOLLEFSON, Nona. "A Comparison of the Item Difficulty and Item Discrimination of Multiple-choice Items Using the «None of the Above» and One Correct Response Options", *Educational and Psychological Measurement*, vol. 47, 1987, p. 377 à 383.

TOUSIGNANT, Robert. *Les principes de la mesure et de l'évaluation des apprentissages*, Saint-Jean-sur-Richelieu, Éditions Préfontaine inc., 1982, 225 p.

WASIK, B.H. et M.D. BOVEN. "Classroom Observational Data : Sources of Inaccuracy and Proposed Solutions", *Behavioral Assessment*, 1980, p. 211-227.

WILLIAMS, Reed G.et Thomas M. HALADYNA, "Logical Operations for Generating Intended Questions (LOGIQ) : A Typology for Higher Level Test Items", *A Technology for Test-item Writing*, Academic Press, Inc., 1982, p. 161 à 186.

WORTHEN, Blaine R. et James R. SANDERS. *Educational Evaluation : Theory and Practice*, Belmont, California Wadsworth Publishing Company, Inc., 1973, 372 p.